



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 15, 2017 – 04:58 am GMT

PDB ID : 2XNX
Title : BC1 FRAGMENT OF STREPTOCOCCAL M1 PROTEIN IN COMPLEX
WITH HUMAN FIBRINOGEN
Authors : Macheboeuf, P.; Y Fu, C.; Zinkernagel, A.S.; Johnson, J.E.; Nizet, V.; Ghosh,
P.
Deposited on : 2010-08-06
Resolution : 3.30 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.9-1692
EDS	:	trunk28620
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	recalc28949

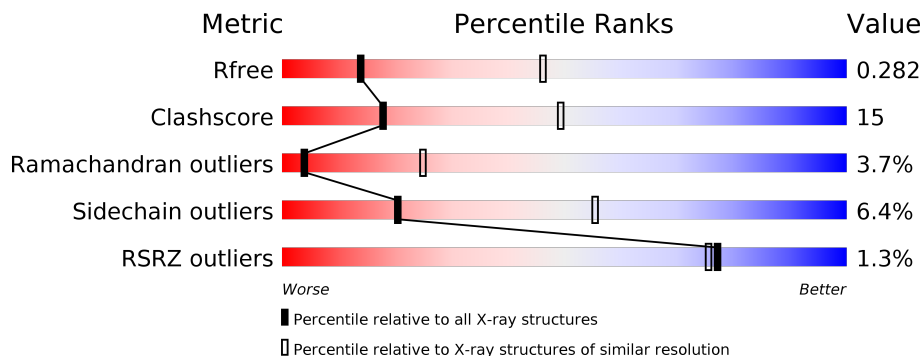
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.30 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	100719	1034 (3.36-3.24)
Clashscore	112137	1100 (3.36-3.24)
Ramachandran outliers	110173	1081 (3.36-3.24)
Sidechain outliers	110143	1080 (3.36-3.24)
RSRZ outliers	101464	1039 (3.36-3.24)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	87	<div> <div>%</div> <div> <div></div> <div>55%</div> <div>25%</div> <div>6%</div> <div>14%</div> </div> </div>
1	D	87	<div> <div></div> <div>55%</div> <div>28%</div> <div>5%</div> <div>13%</div> </div>
1	G	87	<div> <div>%</div> <div> <div></div> <div>56%</div> <div>21%</div> <div>7%</div> <div>16%</div> </div> </div>
1	J	87	<div> <div>%</div> <div> <div></div> <div>51%</div> <div>24%</div> <div>7%</div> <div>17%</div> </div> </div>
2	B	328	<div> <div></div> <div>57%</div> <div>35%</div> <div>• 5%</div> </div>
2	E	328	<div> <div>%</div> <div> <div></div> <div>58%</div> <div>34%</div> <div>• 5%</div> </div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	H	328	<div> <div> <div>2%</div> <div>61%</div> <div>32%</div> <div>5%</div> </div> </div>
2	K	328	<div> <div> <div>%</div> <div>60%</div> <div>32%</div> <div>6%</div> </div> </div>
3	C	319	<div> <div> <div>%</div> <div>65%</div> <div>27%</div> <div>5%</div> </div> </div>
3	F	319	<div> <div> <div>%</div> <div>62%</div> <div>30%</div> <div>5%</div> </div> </div>
3	I	319	<div> <div> <div>3%</div> <div>66%</div> <div>26%</div> <div>6%</div> </div> </div>
3	L	319	<div> <div> <div>2%</div> <div>62%</div> <div>31%</div> <div>.</div> <div>.</div> </div> </div>
4	M	146	<div> <div> <div>%</div> <div>53%</div> <div>14%</div> <div>5%</div> <div>.</div> <div>26%</div> </div> </div>
4	N	146	<div> <div> <div>3%</div> <div>51%</div> <div>16%</div> <div>5%</div> <div>.</div> <div>26%</div> </div> </div>

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 23867 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called FIBRINOGEN ALPHA CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	75	Total	C	N	O	S	0	0	1
			609	377	116	113	3			
1	D	76	Total	C	N	O	S	0	0	0
			628	388	123	114	3			
1	G	73	Total	C	N	O	S	0	0	1
			591	366	113	109	3			
1	J	72	Total	C	N	O	S	0	0	0
			591	366	113	109	3			

- Molecule 2 is a protein called FIBRINOGEN BETA CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	313	Total	C	N	O	S	0	0	1
			2509	1565	443	479	22			
2	E	310	Total	C	N	O	S	0	0	1
			2483	1549	437	475	22			
2	H	313	Total	C	N	O	S	0	0	0
			2517	1571	444	480	22			
2	K	309	Total	C	N	O	S	0	0	1
			2474	1544	435	473	22			

- Molecule 3 is a protein called FIBRINOGEN GAMMA CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	303	Total	C	N	O	S	0	0	1
			2428	1540	408	468	12			
3	F	304	Total	C	N	O	S	0	0	1
			2436	1546	409	469	12			
3	I	300	Total	C	N	O	S	0	0	0
			2409	1530	405	462	12			
3	L	305	Total	C	N	O	S	0	0	1
			2444	1551	410	470	13			

- Molecule 4 is a protein called M PROTEIN.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	M	108	Total	C	N	O	0	0	1
			874	536	153	185			
4	N	108	Total	C	N	O	0	0	1
			874	536	153	185			

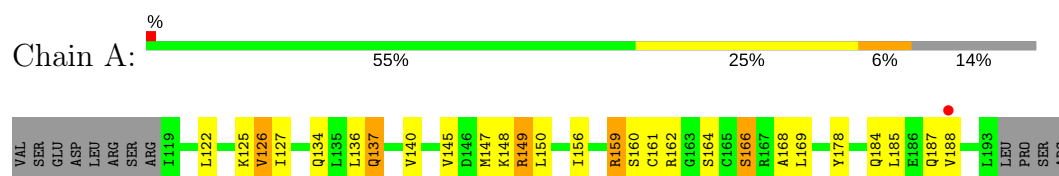
There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
M	126	MET	-	EXPRESSION TAG	UNP Q48WD8
M	127	VAL	-	EXPRESSION TAG	UNP Q48WD8
M	264	LEU	-	EXPRESSION TAG	UNP Q48WD8
M	265	GLU	-	EXPRESSION TAG	UNP Q48WD8
M	266	HIS	-	EXPRESSION TAG	UNP Q48WD8
M	267	HIS	-	EXPRESSION TAG	UNP Q48WD8
M	268	HIS	-	EXPRESSION TAG	UNP Q48WD8
M	269	HIS	-	EXPRESSION TAG	UNP Q48WD8
M	270	HIS	-	EXPRESSION TAG	UNP Q48WD8
M	271	HIS	-	EXPRESSION TAG	UNP Q48WD8
M	195	ALA	THR	CONFLICT	UNP Q48WD8
N	126	MET	-	EXPRESSION TAG	UNP Q48WD8
N	127	VAL	-	EXPRESSION TAG	UNP Q48WD8
N	264	LEU	-	EXPRESSION TAG	UNP Q48WD8
N	265	GLU	-	EXPRESSION TAG	UNP Q48WD8
N	266	HIS	-	EXPRESSION TAG	UNP Q48WD8
N	267	HIS	-	EXPRESSION TAG	UNP Q48WD8
N	268	HIS	-	EXPRESSION TAG	UNP Q48WD8
N	269	HIS	-	EXPRESSION TAG	UNP Q48WD8
N	270	HIS	-	EXPRESSION TAG	UNP Q48WD8
N	271	HIS	-	EXPRESSION TAG	UNP Q48WD8
N	195	ALA	THR	CONFLICT	UNP Q48WD8

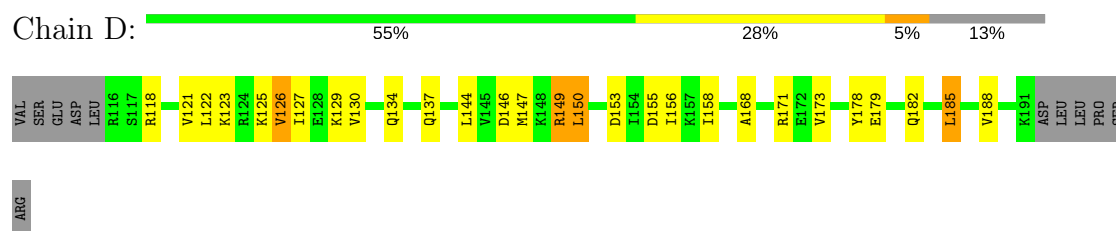
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ($RSRZ > 2$). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

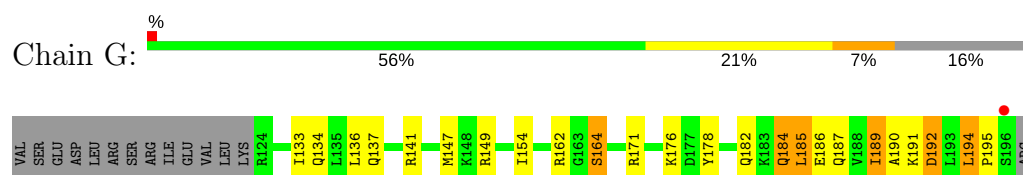
• Molecule 1: FIBRINOGEN ALPHA CHAIN



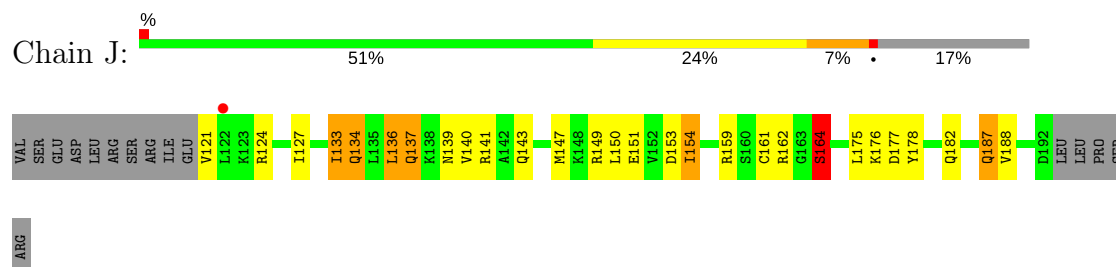
• Molecule 1: FIBRINOGEN ALPHA CHAIN



• Molecule 1: FIBRINOGEN ALPHA CHAIN

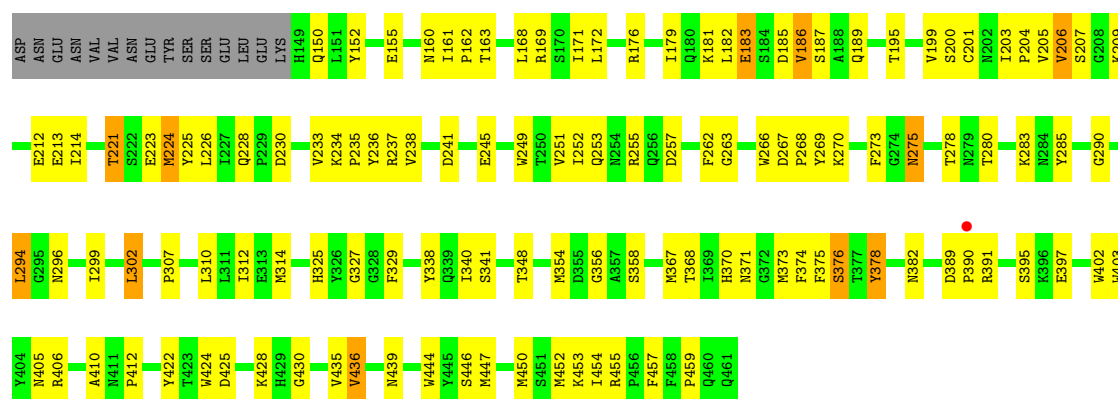


• Molecule 1: FIBRINOGEN ALPHA CHAIN

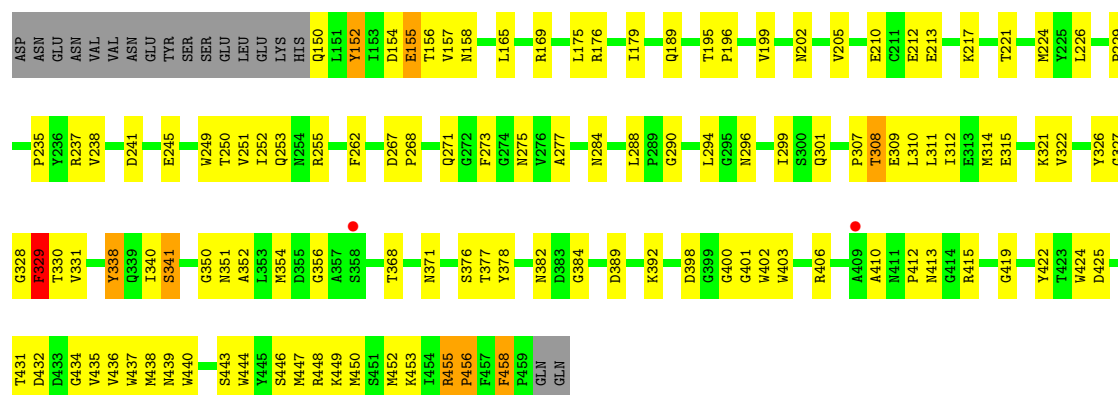


• Molecule 2: FIBRINOGEN BETA CHAIN

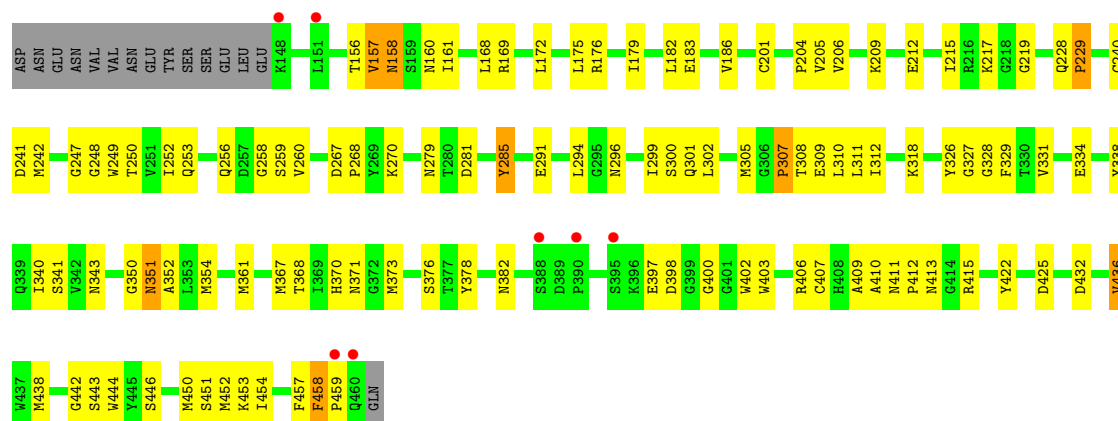




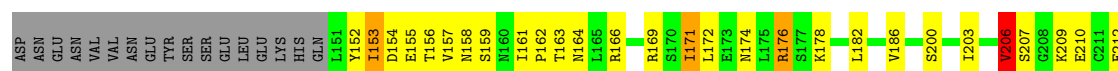
• Molecule 2: FIBRINOGEN BETA CHAIN

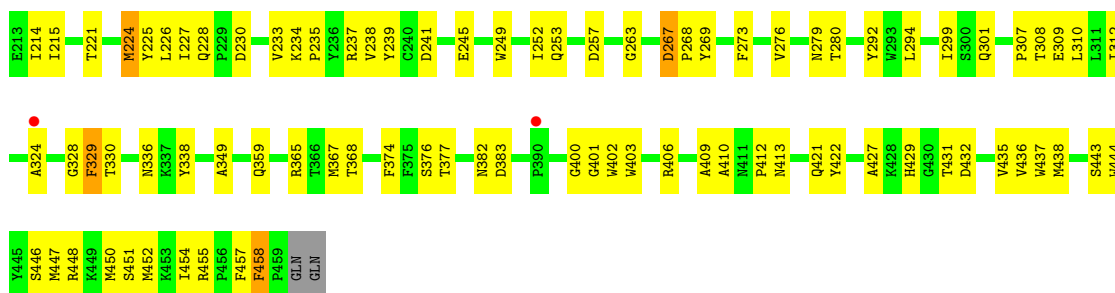


• Molecule 2: FIBRINOGEN BETA CHAIN

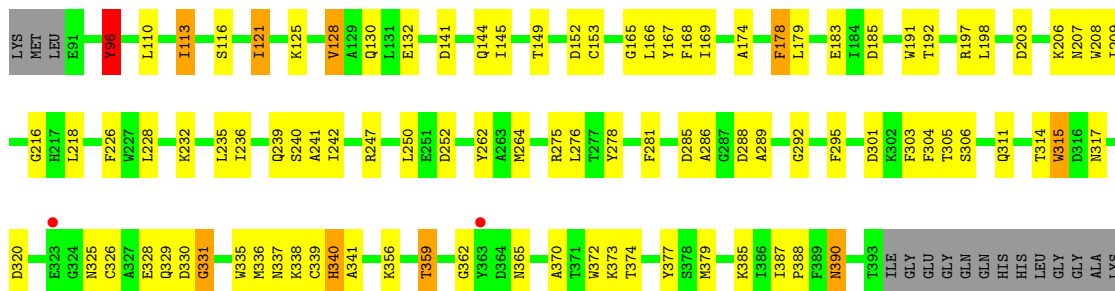


• Molecule 2: FIBRINOGEN BETA CHAIN

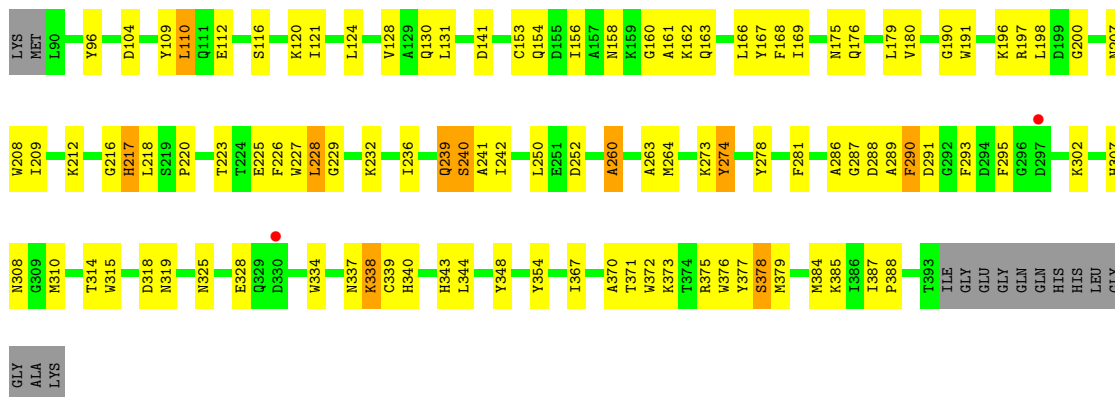




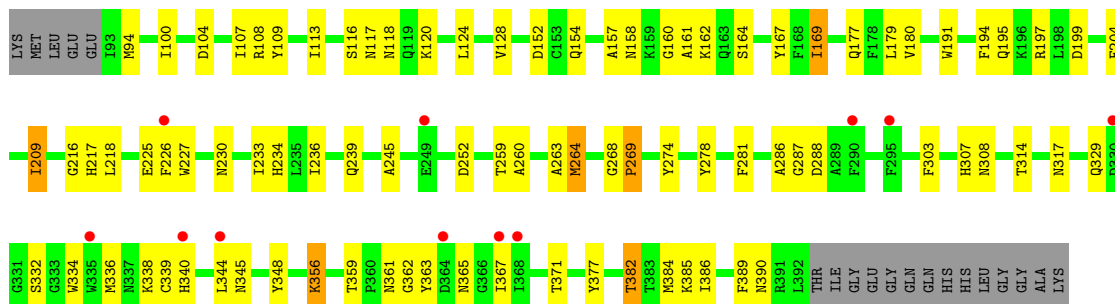
• Molecule 3: FIBRINOGEN GAMMA CHAIN



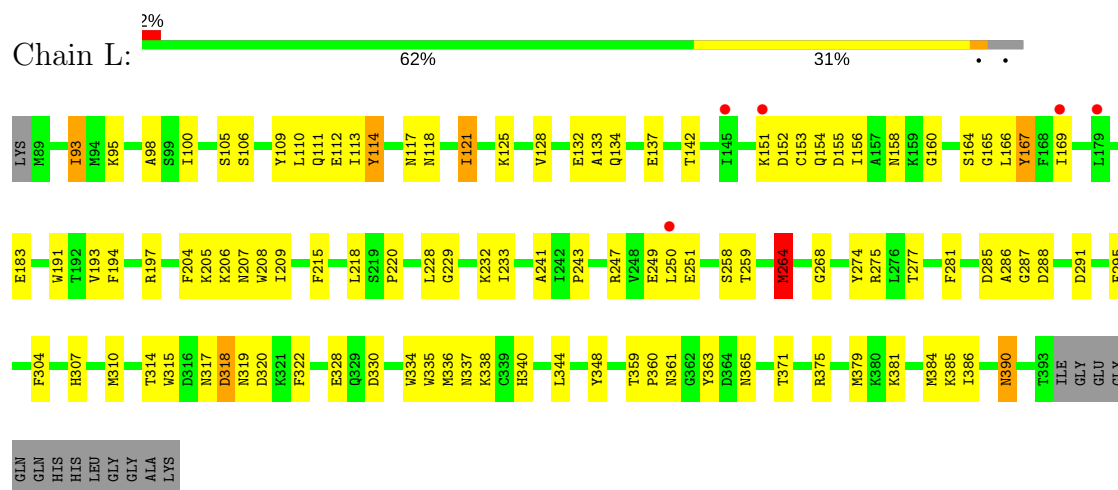
• Molecule 3: FIBRINOGEN GAMMA CHAIN



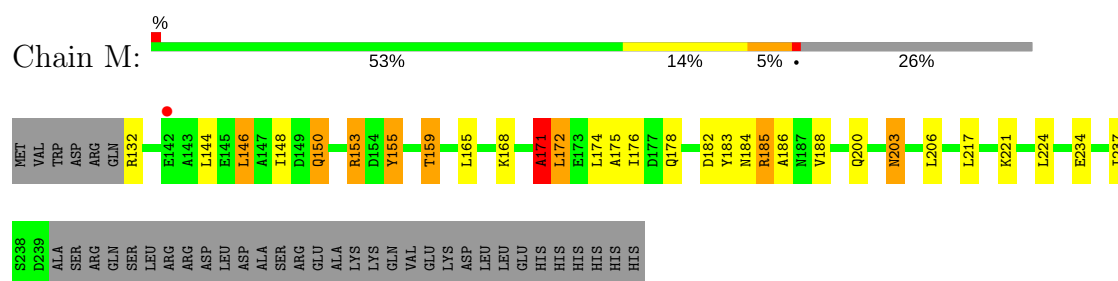
• Molecule 3: FIBRINOGEN GAMMA CHAIN



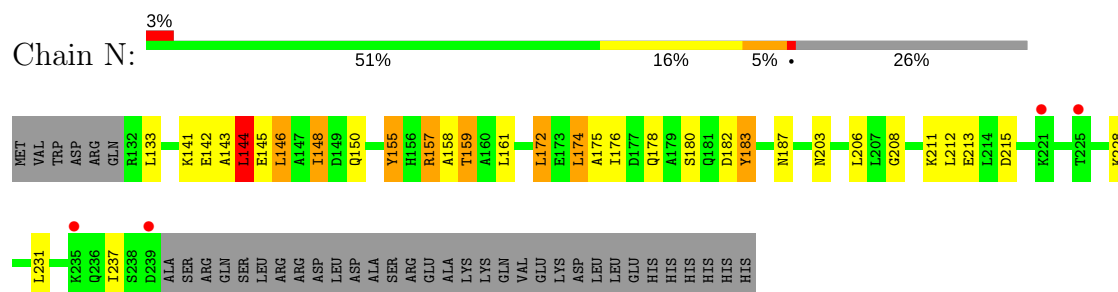
• Molecule 3: FIBRINOGEN GAMMA CHAIN



• Molecule 4: M PROTEIN



• Molecule 4: M PROTEIN



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	112.72Å 216.87Å 140.81Å 90.00° 102.54° 90.00°	Depositor
Resolution (Å)	116.25 – 3.30 116.09 – 3.30	Depositor EDS
% Data completeness (in resolution range)	96.5 (116.25-3.30) 96.5 (116.09-3.30)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.97 (at 3.33Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.277 , 0.325 0.241 , 0.282	Depositor DCC
R_{free} test set	4832 reflections (5.33%)	DCC
Wilson B-factor (Å ²)	80.6	Xtriage
Anisotropy	0.582	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.27 , 58.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.44$, $\langle L^2 \rangle = 0.27$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.88	EDS
Total number of atoms	23867	wwPDB-VP
Average B, all atoms (Å ²)	77.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.38% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# $ Z > 5$	RMSZ	# $ Z > 5$
1	A	0.41	0/610	0.59	0/813
1	D	0.39	0/629	0.57	0/836
1	G	0.41	0/593	0.55	0/792
1	J	0.38	0/592	0.58	0/788
2	B	0.43	0/2573	0.62	0/3478
2	E	0.44	1/2545 (0.0%)	0.62	0/3439
2	H	0.41	0/2581	0.58	0/3487
2	K	0.44	1/2536 (0.0%)	0.59	0/3427
3	C	0.46	0/2494	0.57	0/3372
3	F	0.45	0/2502	0.59	0/3383
3	I	0.41	0/2475	0.53	0/3346
3	L	0.45	0/2510	0.57	0/3393
4	M	0.40	0/876	0.64	1/1169 (0.1%)
4	N	0.39	0/876	0.62	0/1169
All	All	0.43	2/24392 (0.0%)	0.59	1/32892 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	K	458	PHE	C-N	-5.67	1.23	1.34
2	E	458	PHE	C-N	-5.61	1.23	1.34

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	M	171	ALA	C-N-CA	5.07	134.38	121.70

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts ⓘ

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	609	0	649	18	0
1	D	628	0	676	23	0
1	G	591	0	628	16	0
1	J	591	0	632	25	0
2	B	2509	0	2367	98	0
2	E	2483	0	2345	94	0
2	H	2517	0	2380	85	0
2	K	2474	0	2337	79	0
3	C	2428	0	2273	72	0
3	F	2436	0	2284	74	0
3	I	2409	0	2261	56	0
3	L	2444	0	2293	82	0
4	M	874	0	886	32	0
4	N	874	0	886	43	0
All	All	23867	0	22897	712	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 15.

The worst 5 of 712 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:M:171:ALA:HB3	4:M:172:LEU:HB2	1.24	1.14
2:B:169:ARG:HD3	4:N:155:TYR:CE2	1.97	0.97
4:M:171:ALA:HB3	4:M:172:LEU:CB	1.96	0.95
4:N:143:ALA:O	4:N:146:LEU:HB2	1.66	0.93
2:K:412:PRO:HB3	2:K:450:MET:HG2	1.55	0.87

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	73/87 (84%)	53 (73%)	15 (20%)	5 (7%)	1	10
1	D	74/87 (85%)	58 (78%)	14 (19%)	2 (3%)	6	32
1	G	71/87 (82%)	54 (76%)	8 (11%)	9 (13%)	0	2
1	J	70/87 (80%)	52 (74%)	13 (19%)	5 (7%)	1	10
2	B	311/328 (95%)	248 (80%)	52 (17%)	11 (4%)	4	26
2	E	308/328 (94%)	244 (79%)	51 (17%)	13 (4%)	3	22
2	H	311/328 (95%)	255 (82%)	43 (14%)	13 (4%)	3	22
2	K	307/328 (94%)	248 (81%)	47 (15%)	12 (4%)	3	23
3	C	301/319 (94%)	235 (78%)	60 (20%)	6 (2%)	9	39
3	F	302/319 (95%)	244 (81%)	48 (16%)	10 (3%)	4	28
3	I	298/319 (93%)	243 (82%)	48 (16%)	7 (2%)	7	36
3	L	303/319 (95%)	245 (81%)	46 (15%)	12 (4%)	3	23
4	M	106/146 (73%)	93 (88%)	10 (9%)	3 (3%)	6	32
4	N	106/146 (73%)	97 (92%)	7 (7%)	2 (2%)	9	41
All	All	2941/3228 (91%)	2369 (81%)	462 (16%)	110 (4%)	4	25

5 of 110 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	338	TYR
2	E	341	SER
2	E	439	ASN
3	F	240	SER
1	G	192	ASP

5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	69/82 (84%)	64 (93%)	5 (7%)	17	50
1	D	71/82 (87%)	67 (94%)	4 (6%)	25	60
1	G	67/82 (82%)	64 (96%)	3 (4%)	32	67
1	J	67/82 (82%)	59 (88%)	8 (12%)	6	26
2	B	270/286 (94%)	252 (93%)	18 (7%)	19	53
2	E	267/286 (93%)	256 (96%)	11 (4%)	35	69
2	H	271/286 (95%)	257 (95%)	14 (5%)	27	63
2	K	266/286 (93%)	246 (92%)	20 (8%)	16	48
3	C	255/267 (96%)	239 (94%)	16 (6%)	21	56
3	F	256/267 (96%)	242 (94%)	14 (6%)	25	61
3	I	253/267 (95%)	238 (94%)	15 (6%)	23	58
3	L	257/267 (96%)	245 (95%)	12 (5%)	30	66
4	M	94/130 (72%)	83 (88%)	11 (12%)	6	27
4	N	94/130 (72%)	82 (87%)	12 (13%)	5	22
All	All	2557/2800 (91%)	2394 (94%)	163 (6%)	20	55

5 of 163 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	H	161	ILE
3	I	199	ASP
4	M	217	LEU
2	H	206	VAL
2	H	446	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
3	I	337	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	J	182	GLN
2	K	411	ASN
2	H	411	ASN
2	K	336	ASN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled ‘#RSRZ > 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q < 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
1	A	75/87 (86%)	0.24	1 (1%) 77 75	42, 82, 154, 183	0
1	D	76/87 (87%)	0.23	0 100 100	28, 72, 145, 164	0
1	G	73/87 (83%)	0.22	1 (1%) 75 73	36, 83, 157, 208	0
1	J	72/87 (82%)	0.17	1 (1%) 75 73	41, 87, 139, 162	0
2	B	313/328 (95%)	0.21	1 (0%) 93 93	21, 68, 147, 196	0
2	E	310/328 (94%)	0.29	2 (0%) 89 88	18, 60, 121, 187	0
2	H	313/328 (95%)	0.29	7 (2%) 62 60	21, 71, 140, 196	0
2	K	309/328 (94%)	0.19	2 (0%) 89 88	25, 77, 149, 196	0
3	C	303/319 (94%)	0.25	2 (0%) 87 87	15, 55, 105, 180	0
3	F	304/319 (95%)	0.29	2 (0%) 87 87	20, 61, 112, 208	0
3	I	300/319 (94%)	0.31	11 (3%) 42 38	30, 78, 149, 191	0
3	L	305/319 (95%)	0.28	5 (1%) 72 69	23, 69, 122, 187	0
4	M	108/146 (73%)	0.08	1 (0%) 84 83	38, 126, 180, 208	0
4	N	108/146 (73%)	0.21	4 (3%) 42 38	31, 111, 169, 201	0
All	All	2969/3228 (91%)	0.25	40 (1%) 77 75	15, 70, 150, 208	0

The worst 5 of 40 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	H	390	PRO	6.0
2	H	388	SER	5.9
4	N	239	ASP	5.4
2	H	148	LYS	4.9
2	K	390	PRO	4.7

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.